

Weekly Report(Oct. 29th, 2018. 刘一璟)

工作

1. 完整现有方法在LIDC数据集的调参、对比实验
2. 工作时长: 工作日每日8个小时, 周末共10小时, 共50小时.

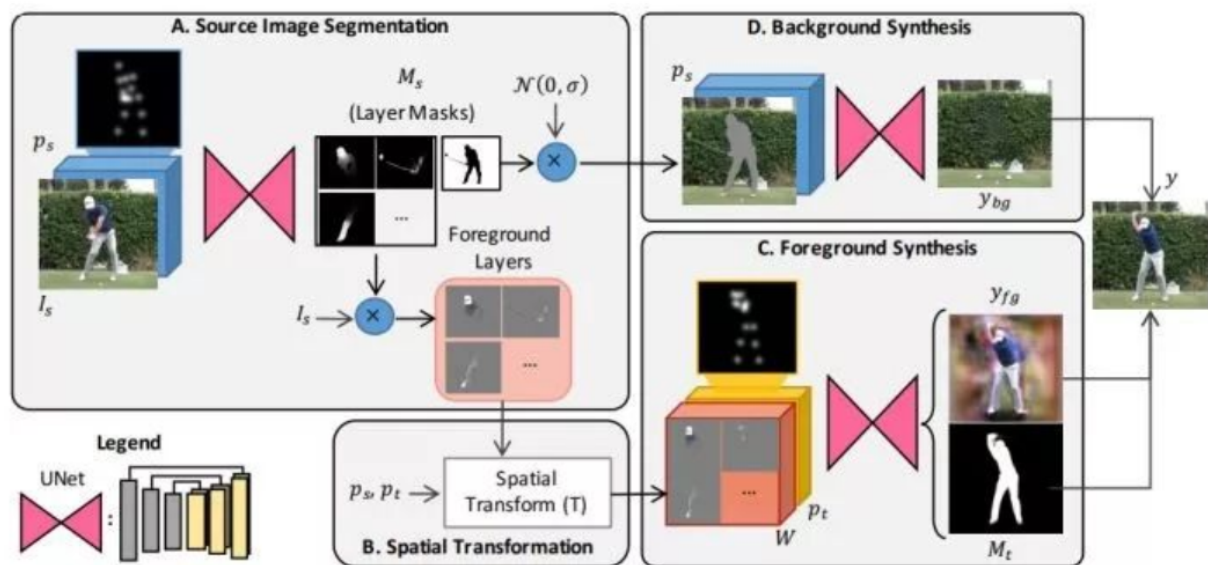
工作进度

项目	进度	截止时间
CVPR 投稿	1.调参部分的实验已经基本完成。对比的实验需要进行交叉验证, 固定好参数在相同网络上进行常用池化与新降采样方法的对比, 还在进行中。	11 月

论文阅读

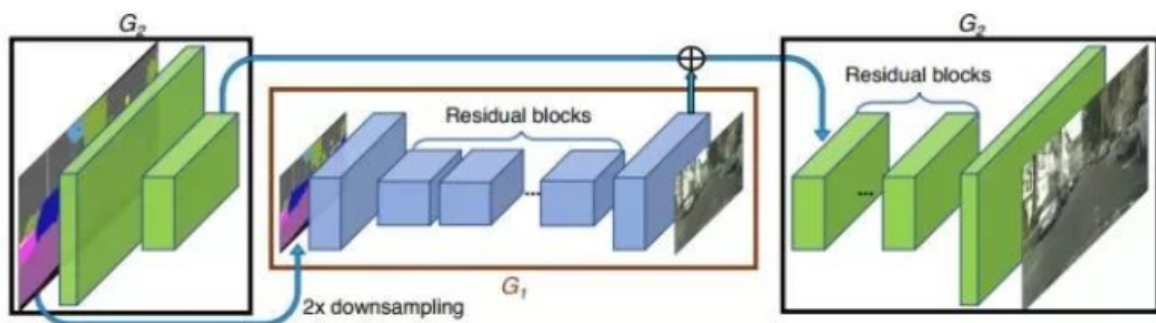
Synthesizing Images of Humans in Unseen Poses

- pose 合成
- 网络的输入是original img, original pose, target pose, 并预设original img和target img背景一样, 人是同一个。首先前后景分离, 然后针对前景(即人), 针对身体的不同部分做细致的segment



High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs

- 解决了GAN生成高分辨率突破的问题，分辨率达到了2048*1024



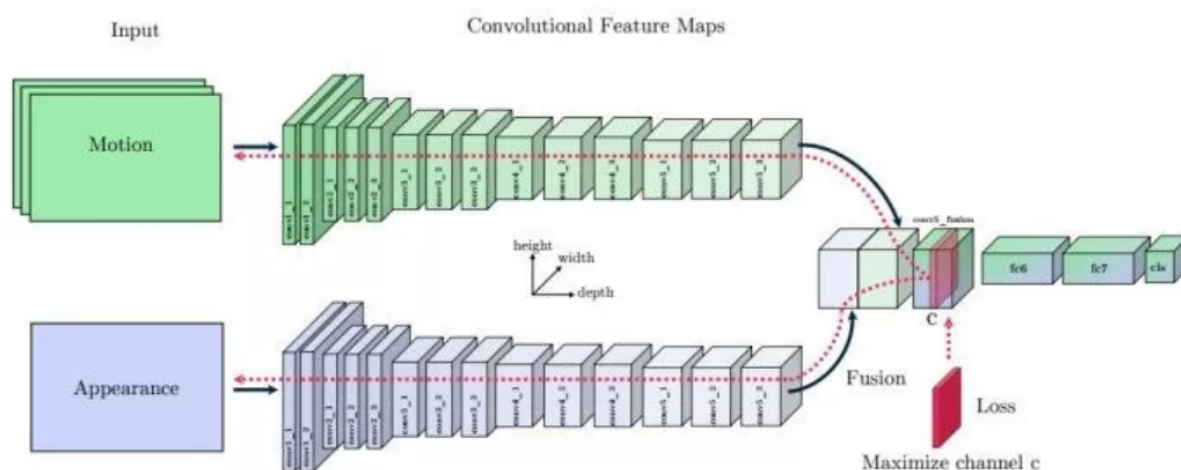
- 这里的discriminator是multi-scale，有着三个针对不同尺寸的D，三个尺寸分别是原尺寸，二分之一，四分之一，仍然是为了解决多尺度问题

$$\min_G \max_{D_1, D_2, D_3} \sum_{k=1,2,3} \mathcal{L}_{\text{GAN}}(G, D_k). \quad (3)$$

- 学习过程使用了边界图片作为输入，提高了效果

What have we learned from deep representations for action recognition?

- 这篇文章就是two-stream模型中间层的可视化方法，换句话说，就是探寻two-stream模型学到了怎样的时空信息



- 生成分为两步： 1. 计算出输入的偏导，将梯度用学习率进行缩放，加入到当前输入
- 提到了两个正则方法： 1.防止过大的值 2. 限制低频信息

$$\mathcal{R}_B(\mathbf{x}) = \begin{cases} N_B(\mathbf{x}) & \forall i, j, k : \sqrt{\sum_d \mathbf{x}(i, j, k, d)^2} \leq B \\ +\infty, & \text{otherwise.} \end{cases}$$

$$\mathcal{R}_{TV}(\mathbf{x}; \kappa, \chi) = \sum_{ijkl} [\kappa ((\nabla_x \mathbf{x})^2 + (\nabla_y \mathbf{x})^2) + \chi (\nabla_t \mathbf{x})^2]$$